The effects of using collaborative assessment with students going abroad: Intercultural competence development, self-understanding, self-confidence, and stages of change

Schnabel, D. B. L.; Kelava, A. ; van de Vijver, Fons

Published in:
Journal of College Student Development

DOI:
10.1353/csd.2016.0000

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

The Effects of Using Collaborative Assessment With Students Going Abroad: Intercultural Competence Development, Self-Understanding, Self-Confidence, and Stages of Change

Deborah B. L. Schnabel, Augustin Kelava, Fons J. R. van de Vijver

Journal of College Student Development, Volume 57, Number 1, January 2016, pp. 79-94 (Article)

Published by Johns Hopkins University Press
DOI: 10.1353/csd.2016.0000

The Effects of Using Collaborative Assessment With Students Going Abroad: Intercultural Competence Development, Self-Understanding, Self-Confidence, and Stages of Change

Deborah B. L. Schnabel Augustin Kelava Fons J. R. van de Vijver

For this study we examined collaborative assessment in counseling 820 German students who were going abroad and who were exposed to the Test to Measure Intercultural Competence (TMIC). A randomized pretest–posttest control group design was used. The control group did not get any test feedback. The remaining groups received written feedback or written plus oral collaborative test feedback. Repeated measures linear mixed effects modeling showed that collaborative test feedback positively influenced students' self-appraisal of their intercultural competence (TMIC-SA); their values on three stages of change; as well as their self-understanding, self-confidence, and perceived benefit from test participation. It is concluded that collaborative assessment and feedback can enhance self-appraised intercultural competence, thereby showing its potential in intercultural training.

Nowadays, global mobility of students is a common practice in most parts of the world. However, development programs that prepare students in higher education for a stay abroad are still exceptional (Straub, Nothnagel, & Weidemann, 2010). The most widespread development activities in the field are culture-specific or cross-cultural awareness group trainings, often offered to students before going abroad (Landis & Bhagat, 1996). These procedures lack focus on the needs, strengths, or weaknesses of the individual (Mendenhall et al., 2004). Personal coaching (or counseling) might be more fruitful. Such a more personalized approach to intercultural competence development is a well-known practice with expatriates (for an overview, see Deardorff, 2010). However, a personalized approach is less often chosen in higher education due to its time and cost intensity (Vulpe, 2004). To our knowledge, there are no evidence-based guidelines that enable counselors to meaningfully interpret and communicate personal results from intercultural competence assessments to foster students' development. The current study addressed this void by investigating the personal benefits of feedback (vs. no feedback) after completing the Test to Measure Intercultural Competence (TMIC; Schnabel, Kelava, van de Vijver, & Seifert, 2015). In the following, we (a) introduce

Deborah B. L. Schnabel is a research fellow and Augustin Kelava is Full Professor, both in the Hector Research Institute of Education Sciences and Psychology at Eberhard Karls Universitaet Tuebingen, Germany. Fons J. R. van de Vijver is Professor of Cross-Cultural Psychology at Tilburg University, The Netherlands, and holds Extraordinary Chairs at North-West University, South Africa, and the University of Queensland, Australia.

the main topics of intercultural competence and collaborative assessment, (b) outline a newly adapted oral collaborative test feedback intervention, and (c) present results of the development effects of collaborative test feedback in a nonclinical setting.

Intercultural Competence

Intercultural competence consists of several facets that enable a person to successfully face unknown challenges while living, working, or studying in different cultures (e.g., Earley & Ang, 2003; Fantini & Tirmizi, 2006; Schnabel, Kelava, Seifert, & Kuhlbrodt, 2015). In contrast to personality traits that are relevant in intercultural interaction (cf. Fantini & Tirmizi, 2006; Kelley & Meyers, 1995; Koester & Olebe, 1988), competences are malleable (Leung, Ang, & Tan, 2014). The most common development method is participation in an intercultural training (e.g., Brislin & Bhawuk, 1999; Landis & Bhagat, 1996). Culture-specific or cross-cultural trainings aim at preparing individuals for successfully handling intercultural situations (Earley, 1987), mainly by building knowledge (Mendenhall et al., 2004). Intercultural coaching focuses on applying this knowledge. The needs and development issues of the client are thus central (Rosinski & Abbott, 2006). In practice, assessment instruments, which measure culture-specific communication and working preferences, are used mostly as part of trainings or coachings but are not used as a developmental intervention per se.

A wide range of instruments used to measure intercultural competence, which differ greatly in their psychometric qualities and operationalization of the construct, has been developed (Gabrenya, Griffith, Moukarzel, Pomerance, & Reid, 2012). Established instruments assess personality traits (e.g., Multicultural Personality Questionnaire; Van Oudenhoven & Van der Zee, 2002), intercultural sensitivity (e.g., Intercultural Development Inventory; Hammer, Bennett, & Wiseman, 2003), or cultural intelligence (e.g., Cultural Intelligence Scale; Van Dyne, Ang, & Koh, 2008). Until recently, almost no instruments addressed malleable abilities. Schnabel, Kelava, Seifert, and Kuhlbrodt (2015) contributed to closing that gap with their recently developed German Test to Measure Intercultural Competence (TMIC). It assesses 17 competence facets that belong to six competence areas (communication, learning, social interaction, self-management, self-knowledge, and building synergies). Three studies showed that the TMIC has favorable psychometric characteristics (Schnabel & Kelava, 2013; Schnabel, Kelava, Seifert, & Kuhlbrodt, 2015; Schnabel, Kelava, van de Vijver, & Seifert, 2015), which underscore the potential of the TMIC to serve as a basis for training and intervention.

Collaborative Assessment

Collaborative assessment (Fischer, 1994), also called therapeutic assessment (Finn, 1996, 2007), is a highly individualized approach to using psychological tests in counseling. Test results are interpreted in light of the personal experiences and situation of a client who works together with the assessor to increase the unique benefit of the assessment process (Finn & Tonsager, 1997; Fischer, 2000). Clients become coassessors who share their opinions in an open and trustful dialogue (Craddick, 1975). Therefore, collaborative assessment goes beyond pure information gathering (Finn & Tonsager, 1997). Waiswol (1995) stated that therapeutic assessment, with its transformative nature, functions like a brief intervention (Miller & Rollnick, 2002). Three basic human motives are regulating the individual's transformation in such an intervention (Finn & Kamphuis, 2006):

- Self-verification (Swann & Read, 1981), which is addressed in collaborative assessment through confirmation of the client's own view of the personal self;
- Self-enhancement (Sedikides & Gregg, 2008), which is fulfilled when the client feels valued; and
- Self-efficacy-self-discovery (Bandura, 1994), which is satisfied when the client learns novel aspects about him or herself and about—so far—unsolved problems (Finn & Kamphuis, 2006).

The process of change induced by collaborative assessment can be explained by Prochaska's transtheoretical model of stages of change (TTM; see DiClemente & Prochaska, 1998; Prochaska, DiClemente, & Norcross, 1992). The TTM assumes that behavioral change is an intentional process consisting of multiple phases. The process starts with a precontemplation stage, in which an individual denies any problematic behavior or need to change. Successively, the individual starts to accept, plan, and take actions to change behavior (DiClemente & Prochaska, 1998). Finn and Kamphuis (2006) assumed that collaborative assessment positively influences the localization of an individual in one of the early three stages of change. However, this has not yet been empirically investigated.

Therapeutic Outcome of Collaborative Assessment and Feedback Interventions

Poston and Hanson (2010) conducted a meta-analysis including 17 studies dealing with psychological assessment as a therapeutic intervention. Treatment group means were significantly higher than were the means in the reference group in 66% of the comparisons, $d = 0.423$, 95% CI [0.321, 0.525]. Furthermore, moderate treatment group effects have been found for therapy outcomes, $d = 0.367$, 95% CI [0.256, 0.478], such as increased self-esteem, greater self-awareness and self-understanding, higher motivation to seek mental health therapy, and more satisfaction with the feedback (Allen, Montgomery, Tubman, Frazier, & Escovar, 2003; Finn & Tonsager, 1992; Newman & Greenway, 1997).

To our knowledge, no study has addressed the effects of the combination of competence assessment and feedback on intercultural competence thus far. However, evidence for the impact of test feedback on learning can be drawn from educational research. Collaborative assessment can be integrated into formative assessment that concentrates on improvement. The information-gathering approach functions as an equivalent to summative assessment, which focuses on describing the current state of the assessee (Crooks, 2001). The learning success factor in formative assessment and collaborative assessment is adequate feedback (Black & Wiliam, 1998; Crooks, 2001; Finn & Tonsager, 1997; Fischer, 2000). There are several meta-analyses highlighting the power of feedback in education (e.g., Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Moin, 1986). Probably the most authoritative study was conducted by Kluger and DeNisi (1996), who found that feedback interventions increased students' performance in various areas, $d = 0.41$.

Performance enhancement through feedback is attained mainly by addressing the gap between the current level of competence and the desired level (Hattie & Timperley, 2007; Miller & Rollnick, 2002; Sadler, 1989). This is achieved both by collaborative processing of information and by working with and toward goals (see Balcazar, Hopkins, & Suarez, 1986; Finn, 1996; Hattie, 1999; Locke & Latham, 1984). Moreover, change is facilitated by feedback that focuses on improvement instead of mirroring the current level of competence (Crooks, 2001). As mentioned before, this is also how collaborative

assessment differs from information-gathering approaches to psychological tests.

Concerning the mode of providing assessment feedback, Folds and Gazda (1966) showed that test participants who received individual feedback were more satisfied with the test interpretation procedure than were those getting a written report. Their three feedback groups also significantly changed in self-concept, but there was no interaction between the degree of change and the type of test interpretation. Holmes (1964) showed that feedback by a counselor increased the perceived benefit of the test results in students when compared to written feedback via mail. We found it interesting that testing without any feedback seems as meaningless as receiving no treatment at all (Poston & Hanson, 2010). In a more recent study, Lance and Krishnamurthy (2003) examined the effect of a combined written and oral feedback approach on client satisfaction. According to their results, oral feedback alone does not lead to a greater satisfaction than written feedback. This makes the combined oral and written feedback most preferable.

This Study

In accordance with the findings reviewed above, the following hypotheses are tested:

Hypothesis 1: In comparison to no or written feedback alone, combined written and oral collaborative test feedback positively influences the self-appraisal of students' intercultural competence.

Hypothesis 2a–c: Students receiving combined written and oral collaborative test feedback show (a) a better self-understanding of their intercultural competence, (b) a higher self-confidence of mastering intercultural situations, and (c) a greater perceived benefit of taking part in the TMIC-SA after treatment than do individuals without any or with written feedback alone.

Hypothesis 3a–c: Combined written and oral collaborative test feedback has a positive effect on an individual's stage of change. Specifically, subjects in the intervention group have higher means on the (a) contemplation and (b) action stage but (c) lower means on the precontemplation stage (Prochaska & DiClemente, 1982) following feedback than do participants in the comparison or control group.

Hypothesis 4a–b: (a) Students in the combined written and oral collaborative test feedback group report higher satisfaction with the feedback than those in the written feedback group; (b) the more satisfied participants are with the oral part of the collaborative test feedback, the higher is their gain in intercultural competence after the session.

This study adds to the existing research by strengthening the external validity of collaborative assessment in highlighting its effects for students' self-related and intercultural competence development in a nonclinical setting. Additionally, we enrich the utility (see Hayes, Nelson, & Jarrett, 1987) of (intercultural) competence assessments in student counseling, coaching, and training practice.

METHOD

Sampling Procedures

We recruited participants from a German university. We cooperated with the Department for International and External Affairs, which organizes in- and outbound activities of the university as well as intercultural development activities. The sample consisted of students who were already accepted for a year-abroad program in Europe, Asia, or America, which were to start approximately four months after we began the study (April 2013). We randomly selected 88 outgoing students who were invited to take part in an advanced preparation program that consisted of taking the TMIC-SA and receiving oral collaborative test feedback

on the one hand and attending an intercultural training—realized after the posttest and hence not part of the study—on the other hand. Finally, 73 of the invited 88 outgoing students in the treatment group participated in the pretest. Out of the remaining 794 subjects, we randomly selected students for the comparison and control groups, both of which initially consisted of 397 potential study participants. A total of 351 students in the no feedback group and 396 students in the written feedback condition took part in the pretest. When comparing the number of participants between the pretest and the posttest (see Table 1 for number of participants in the posttest), the dropout rates for the three groups (no feedback group, written feedback group, and written plus oral collaborative test feedback group) were 56%, 89%, and 55%, respectively. A detailed analysis of potential systematic effects in relation to those drop-outs is given at the beginning of the Results section.

The invitation for the follow-up survey was sent 2 weeks after the first session. All participants had 2 weeks to fill in the follow-up survey. Two reminders were sent in between. After the pretest, participants could sign up for a lottery with the possibility to win one of three Amazon vouchers (two for 25 euro and one for 50 euro) or an Apple Ipod Nano.

Sample Size, Structure, and Power

Altogether, 820 students—480 females (58%), 327 males (40%), and 13 not indicating their sex (2%)—with an average age of 23.37 years (SD = 3.89) took part in the pretest. In the follow-up study, 233 students participated. These 133 men (57%) and 100 women (43%) were 23.47 years old (SD = 2.89) on average. The age and sex for all three subgroups as well as results concerning education and intercultural experiences, which were similar across the three groups, are shown in Table 1.

Measures

Test to Measure Intercultural Competence (TMIC). In this study, we used the Self-Appraisal scale of the Test to Measure Intercultural Competence (TMIC-SA; Schnabel, Kelava, Seifert, & Kuhlbrodt, 2015). TMIC-SA has 75 items (with responses on a 6-point Likert-type scale ranging from *does not apply at all* to *fully applies*) to measure an individual's self-concept concerning the six intercultural competence areas mentioned before. Example items for each of the six intercultural competence areas are as follows (all example items have been translated from German to English for this article; however, these English items have not been validated): "The way I address something depends on the person I am talking to" (communication), "When planning a trip abroad I use various sources of information" (learning), "When I join a group for the first time I quickly build relationships with other group members" (social interaction), "When I plan something I usually then go on to achieve my aim" (self-management), "I am good at mediating between people with conflicting interests" (creating synergies), and "I make an effort to understand to what extent my behavior is shaped by culture" (self-knowledge).

Exploratory structural equation modeling (Asparouhov & Muthén, 2009) revealed a good model fit for the 17 competence facets, $\chi^2(1636, N = 641) = 2579.85$, $p < .001$; $\chi^2/df = 1.58$; RMSEA = .031, RMSEA 90% CI [.029–.033]; SRMR = .017; CFI = .955; TLI = .927. Cronbach's alpha of the TMIC-SA indicated a high internal consistency of the scale ($\alpha = .96$; Schnabel, Kelava, Seifert, & Kuhlbrodt, 2015). Additionally, metric invariance was established for a German and Brazilian short version of the TMIC (TMIC-S; Schnabel, Kelava, van de Vijver, & Seifert, 2015).

TABLE 1.

Sociodemographic Characteristics and Intercultural Experiences Divided by Subgroups (Percentages in Parentheses)

Characteristics	NFB (*n* = 156)	WFB (*n* = 44)	WOFB (*n* = 33)
Age	*M* = 23.55, *SD* = 3.02	*M* = 22.93, *SD* = 2.59	*M* = 23.82, *SD* = 2.60
Sex			
Male	93 (60)	25 (57)	15 (46)
Female	63 (40)	19 (43)	18 (55)
Target Degree			
Bachelor	55 (35)	19 (43)	15 (46)
Master	51 (33)	15 (34)	11 (33)
Diploma	3 (2)	2 (5)	2 (6)
PhD	18 (11)	4 (9)	3 (9)
Other	29 (19)	4 (9)	2 (6)
Study Area			
Humanities	70 (45)	17 (39)	8 (24)
Sciences	44 (28)	13 (29)	6 (18)
Engineering	28 (18)	10 (23)	16 (49)
Other	14 (9)	4 (9)	2 (9)
Intercultural Training Experiences			
Yes	47 (30)	10 (23)	9 (27)
No	109 (70)	34 (77)	24 (73)
Intercultural Involvement			
Yes	94 (60)	24 (55)	19 (58)
No	62 (40)	20 (46)	14 (42)
Previous Experiences Abroad			
Internship	52 (15)	12 (14)	12 (17)
Project	49 (14)	13 (15)	9 (13)
Studying	75 (21)	17 (20)	11 (15)
Vacation	132 (38)	27 (43)	30 (42)
Work	42 (12)	7 (8)	9 (13)

Note. NFB = No feedback group; WFB = Written feedback group; WOFB = Written plus oral collaborative test feedback group. The difference in age between the three groups was not significant, $F(2, 228) = 1.05$, $p = .350$.

Stages of Change. The few German questionnaires that measure the TTM are related to drug or alcohol abuse. Therefore, items had to be adapted. We started with the original items by Fecht, Heidenreich, Hoyer, Lauterbach, and Schneider (1998); Hoyer, Heidenreich, Fecht, Lauterbach, and Schneider (2003); and Hannöver, Rumpf, Meyer, Hapke, and John (2001). It was not possible to reach maintenance, the fourth phase, through the applied written and oral collaborative test feedback intervention. Therefore, we restricted the measure to three stages: precontemplation (e.g., "I guess I have weaknesses in the area

TABLE 2.
Means, Standard Deviations, and Cronbach's alpha for all Dependent Variables

	Total Sample Pretest (*N* = 820)				Remaining Sample Posttest (*N* = 233)			
	NFB (*n* = 351)	WFB (*n* = 396)	WOFB (*n* = 73)	Total (*N* = 820)	NFB (*n* = 156)	WFB (*n* = 44)	WOFB (*n* = 33)	Total (*N* = 233)
Dependent Variables	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	**α**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	**α**
TMIC-SA T0	4.41 (.46)	4.41 (.47)	4.33 (.55)	.95	4.43 (.43)	4.22 (.52)	4.42 (.46)	.95
TMIC-SA T1					4.38 (.45)	4.19 (.63)	4.65 (.42)	
Precontemplation T0	2.72 (.81)	2.74 (.86)	2.67 (.93)	.66	2.68 (.79)	2.76 (1.03)	2.47 (.75)	.70
Precontemplation T1					2.83 (.78)	2.83 (.99)	2.12 (.68)	
Contemplation T0	3.30 (.87)	3.35 (.91)	3.97 (.77)	.75	3.38 (.87)	3.35 (.81)	3.92 (.63)	.73
Contemplation T1					3.40 (.90)	3.38 (.95)	4.27 (.79)	
Action T0	3.33 (.93)	3.29 (.87)	3.93 (.88)	.70	3.38 (.94)	2.93 (.93)	3.89 (.74)	.74
Action T1					3.30 (.85)	3.26 (.92)	4.59 (.65)	
Perceived Benefit T0	3.44 (.79)	3.56 (.98)	3.70 (.80)	.89	3.50 (.81)	3.70 (.98)	3.80 (.74)	.87
Perceived Benefit T1					3.69 (.99)	3.47 (.98)	4.88 (.90)	
Self-Confidence T0	4.37 (.76)	4.30 (.73)	4.38 (.74)	.79	4.39 (.72)	4.08 (.60)	4.39 (.68)	.78
Self-Confidence T1					4.36 (.73)	4.24 (.68)	4.75 (.61)	
Self-Understanding T0	3.66 (.89)	3.60 (.78)	3.55 (.86)	.89	3.65 (.88)	3.30 (.69)	3.44 (.77)	.90
Self-Understanding T1					3.59 (.82)	3.66 (.84)	4.77 (.67)	

Note. NFB = no feedback group; WFB = written feedback group; WOFB = written plus oral collaborative test feedback group. T0 = pretest; T1 = posttest.

TABLE 3.
Multivariate Randomized Pretest–Posttest Control Group Design

	Group	Pretest	Treatment	Posttest
R	G1 ($n1$ = 156)	$\bar{Y}1$–7, pre	No Feedback (×0)	$\bar{Y}1$–7, post
	G2 ($n2$ = 44)	$\bar{Y}1$–8, pre	Written Feedback (×1)	$\bar{Y}1$–8, post
	G3 ($n3$ = 33)	$\bar{Y}1$–8, pre	Written Plus Oral Collaborative Test Feedback (×2)	$\bar{Y}1$–9, post

Note. *R* = randomized; G = group; $\bar{Y}$ = means of dependent variables; $\bar{Y}1$ = TMIC-SA total score, $\bar{Y}2$ = Precontemplation Stage, $\bar{Y}3$ = Contemplation Stage, $\bar{Y}4$ = Action Stage, $\bar{Y}5$ = Self-Understanding, $\bar{Y}6$ = Self-Confidence, $\bar{Y}7$ = perceived benefit from TMIC participation, $\bar{Y}8$ = general satisfaction with the written feedback, $\bar{Y}9$ = specific satisfaction with the oral session of the collaborative test feedback SHORT.

of intercultural competence, but there is nothing that I really have to change about it"), contemplation (e.g., "I should inform myself about how I could possibly foster my intercultural competence"), and action (e.g., "I work hard on changing myself"). Four items with responses given on a 6-point Likert-type scale measured each stage.

Treatment Benefit (Outcome) Variables. We used previous therapeutic assessment evaluation studies (e.g., Allen et al., 2003; Finn & Tonsager, 1992; Newman & Greenway, 1997) to identify domains in which the treatment could have an effect. We adapted items from the well-established Rosenberg Self-Esteem Scale (RSE; Rosenberg, 1979), the Self-Liking/Self-Competence Scale (SLCS-R; Tafarodi & Swann, 1995, 2001), and a subscale of the Assessment Questionnaire–2 (Finn & Tonsager, 1992, 1997), the New Self-Awareness/Understanding so as to focus items on intercultural competence. Specifically, we measured self-understanding with six items (e.g., "I know what is important when interacting with people from other cultures"), self-confidence with four items (e.g., "I feel competent enough to deal with problems, which are arising from working together with people from other cultures"), and perceived benefit from taking part in the test with five items (e.g., "In taking part in the present study, I learned something about myself.").

As was done in other collaborative assessment studies, we independently developed the items assessing the degree of satisfaction with the feedback. We measured general satisfaction in the written and oral collaborative test feedback group with seven items (e.g., "Now, that I got feedback concerning my intercultural competences, I know, which steps I have to take in order to improve"). The scale measuring satisfaction with the oral part of the collaborative test feedback comprised five items (e.g., "My true self was well reflected in the feedback talk"). For all items a 6-point Likert-type scale was used. Psychometric properties of each scale are presented in Table 2. The German items are available from the contact author.

Research Design and Experimental Intervention

The study employed a randomized pretest–posttest control group design with a treatment, a comparison, and a control group. The control group did not get any test feedback or other treatment. The comparison group received written feedback, and the treatment group written plus oral collaborative test feedback. Variables of interest were the total score of intercultural competence measured with the TMIC-SA (Schnabel, Kelava, Seifert, &

Kuhlbrodt, 2015) as well as students' stages of change, self-understanding, self-confidence, perceived benefit from test participation, and satisfaction with the assessment and feedback process (e.g., Allen et al., 2003; Finn & Tonsager, 1992; Hannöver et al., 2001). An overview of groups, variables, and the pretest–posttest design is given in Table 3.

We administered the written feedback report immediately after a person had finished the survey. A brief description of all 17 TMIC facets was presented along with the personal score and its relative position in a reference group of 641 German university students (Schnabel, Kelava, Seifert, & Kuhlbrodt, 2015). Thus, each participant could read from the report whether his or her values in the 17 competence facets and in the overall TMIC-SA scale were below, on, or above average. All students could keep the report.

The oral part of the collaborative test feedback was administered in a telephone interview. One assessor conducted all feedback sessions. She followed a highly structured feedback guideline and used a language free of jargon (Finn, 1996; Mosak & Gushurst, 1972). The maximum duration of a session was 1 hour. Participants received a graphical representation of their results to prepare for the phone call. The graphic was also used during the session to discuss results. The collaborative feedback guideline was adapted partly from the manual for using the Minnesota Multiphasic Personality Inventory (MMPI–2; Hathaway & McKinley, 1942) as a therapeutic intervention (Finn, 1996; Fischer & Finn, 2008) and contained important aspects of motivational interviewing techniques (Miller & Rollnick, 2002). Moreover, we followed the general recommendations for providing psychological test feedback (e.g., American Psychological Association, 1986, 2010). In the following, we briefly outline our intervention, named SHORT, an acronym that stands for the key steps during the oral session of the collaborative test feedback:

1. Start and How We Proceed: Establishing the general framework of the oral session of the collaborative test feedback;
2. Orientation: Clarifying roles, asking to share previous test experiences, dealing with assessment questions, talking about the assessees's plans abroad, explaining the TMIC;
3. Reflection: Discussing results starting with those in line with the assessee's self-concept; and
4. Targets: Prompting to set major development goals and answering final questions.

RESULTS

We first report the analysis of the drop-outs, which includes a comparison of the 233 students who took part in the study twice and the 587 students who did not participate in the follow-up study. In the second part, we review the psychometric properties of the administered scales. In the third part, we evaluate the treatment effects: the effects of the three feedback types on the students' intercultural competence, on the therapeutic benefit variables, and on the stages of change are addressed. We discuss each of the three in the same way. First of all, we compare the values in the dependent variables for the three feedback groups in the pretest, using the whole sample of 820 students. In the next step, we analyze the change of the dependent variables across time. In addition, we analyze test score differences between the three feedback groups in the posttest. These analyses are restricted to students who participated in both parts of the study twice (n = 233). The results section closes with an evaluation of the oral part of the collaborative test feedback (SHORT). Taking the following multiple mean comparisons

into consideration, we decided to employ a Bonferroni-Holm correction (Holm, 1979) on all group comparisons in the results section. Accordingly, results are handled as significant as soon as they reach a $p < .01$ level. All means and standard deviations are displayed in Table 2.

Analysis of Drop-outs

We examined if those individuals who decided not to take part in the second assessment ($n = 587$) scored different initially on the TMIC-SA or on the scales measuring intercultural self-understanding, intercultural self-confidence, and the three stages of change in a multivariate analysis of variance (MANOVA). As more women than men participated in the pretest and more men than women participated in the posttest, we added the variable sex as a second independent variable to our MANOVA. The multivariate effect of group (participated a second time vs. did not participate a second time) was not significant, Wilks's $\Lambda = .98$, $F(7, 784) = 2.66$, $p = .010$, $\eta^2 = .02$. Additionally, no univariate group test was significant: TMIC-SA total score, $F(1, 793) = 4.02$, $MSE = 0.85$, $p = .045$; intercultural self-understanding, $F(1, 793) = 2.69$, $MSE = 1.82$, $p = .101$; intercultural self-confidence, $F(1, 793) = 0.00$, $MSE = 0.00$, $p = .987$; precontemplation stage, $F(1, 793) = 2.08$, $MSE = 1.38$, $p = .150$; contemplation stage, $F(1, 793) = 2.53$, $MSE = 1.97$, $p = .112$; action stage, $F(1, 793) = 0.06$, $MSE = 0.05$, $p = .802$. Moreover, the multivariate sex × group interaction was not significant, Wilks's $\Lambda = .99$, $F(7, 784) = 1.26$, $p = .010$, $\eta^2 = .02$. Because of the small effect size of the multivariate group effect as well as the absence of significance for both the univariate tests and the multivariate interaction Sex × Group effect, we concluded that the dropout was not selective.

Internal Consistencies of the Scales

The test–retest reliability (after 2 weeks) in the no feedback group was high ($r = .89$, $p < .001$). As shown in Table 2, Cronbach's alpha values for all used scales were satisfactory.

Total Score of the TMIC-SA

An analysis of variance revealed no significant difference in the TMIC-SA total score between the three experimental groups in the pretest, $F(2, 815) = 1.03$, $MSE = 0.23$, $p = .357$. The repeated measures linear mixed model resulted in a nonsignificant main effect of time for the TMIC-SA total score, $F(1, 228) = 3.92$, $MSE = 0.81$, $p = .021$, $\eta^2 = .03$. However, the interaction effect was significant, which shows that the change was dependent on the type of feedback a person received, $F(1, 228) = 17.65$, $MSE = 0.49$, $p < .001$, $\eta^2 = .14$. Although the TMIC-SA total score did not differ across time for individuals in the written feedback group, $\Delta M = 0.003$, 95% CI [–0.012, 0.006], TMIC-SA total score significantly increased for the written plus oral collaborative test feedback group, $\Delta M = 0.224$, 95% CI [0.215, 0.234]. This supports hypothesis 1. The TMIC-SA total score slightly decreased in the no feedback group over time, $\Delta M = –0.045$, 95% CI [–0.050, –0.041].

Moreover, the effect of feedback group on intercultural competence was significant for the posttest, $F(2, 227) = 8.05$, $MSE = 1.89$, $p < .001$, $\eta^2 = .07$. We assessed group differences with the help of a priori defined contrasts. They showed that the intervention group differed from the no feedback as well as the written feedback group for the posttest, $t(227) = –3.79$, $p < .001$, $d = –0.50$.

Therapeutic Benefit Variables

For the pretest, there were no differences between the three experimental groups concerning perceived benefit from TMIC participation, $F(2, 803) = 2.29$, $MSE = 2.29$, $p = .057$; intercultural self-confidence, $F(2, 809) = 0.84$, $MSE = 0.46$, $p = .434$; and

intercultural self-understanding, $F(2, 809) = 0.72$, $MSE = 0.50$, $p = .488$.

Three repeated measures linear mixed models examined the therapeutic effects of a TMIC assessment in relation to the received feedback. All three therapeutic benefit variables significantly changed over time. For perceived benefit from TMIC participation, $F(1, 226) = 21.82$, $MSE = 7.52$, $p < .001$, $\eta^2 = .09$, and intercultural self-confidence, $F(1, 226) = 10.77$, $MSE = 1.47$, $p < .01$, $\eta^2 = .05$, the effect was moderate and for self-understanding it was large, $F(1, 226) = 100.29$, $MSE = 19.10$, $p < .001$, $\eta^2 = .31$. As shown by three sizable interaction effects (see Table 4), the direction of change was dependent on the feedback type. Participants who received no or written plus oral collaborative test feedback perceived the benefit after taking part in the TMIC-SA a second time as higher than at the first time, $\Delta M = 0.187$, 95% CI [0.164, 0.210] and $\Delta M = 1.079$, 95% CI [1.029, 1.128], respectively. Moreover, the perceived benefit from taking part in the assessment process decreased after individuals got written feedback, $\Delta M = –0.297$, 95% CI [–0.343, –0.252]. Intercultural self-confidence beliefs were lower in the posttest for the no feedback group, $\Delta M = –0.037$, 95% CI [–0.039, –0.035], and higher for the written feedback and written plus oral collaborative test feedback groups, $\Delta M = 0.109$, 95% CI [0.105, 0.114] and $\Delta M = 0.356$, 95% CI [0.351, 0.361], respectively. We found the largest interaction effect for intercultural self-understanding with individuals having the highest increase in intercultural self-understanding after the written plus oral collaborative test feedback, $\Delta M = 1.324$, 95% CI [1.312, 1.336]. Intercultural self-understanding also grew as a consequence of written feedback, $\Delta M = 0.286$, 95% CI [0.275, 0.298]. If no feedback at all was provided after the pretest, intercultural self-understanding values were

TABLE 4.

F Values, Effect Sizes, Mean Differences, and Confidence Intervals for the Interaction Effect Time × Feedback

	Total ($N = 233$)	NFB ($n = 156$)		WFB ($n = 44$)		WOFB ($n = 33$)	
Dependent Variables	**Interaction Time × Feedback**	**ΔM**	**95% CI**	**ΔM**	**95% CI**	**ΔM**	**95% CI**
Perceived Benefit From TMIC	$F = 25.31$, $MSE = 8.73$, $\eta2 = .18$	0.187	[0.164, 0.210]	–0.297	[–0.343, –0.252]	1.079	[1.029, 1.128]
Self-Confidence	$F = 8.05$, $MSE = 1.10$, $\eta2 = .07$	–0.037	[–0.039, –0.035]	0.109	[0.105, 0.114]	0.356	[0.351, 0.361]
Self-Understanding	$F = 69.55$, $MSE = 13.24$, $\eta2 = .38$	–0.066	[–0.072, –0.061]	0.286	[0.275, 0.298]	1.324	[1.312, 1.336]
Precontemplation Stage	$F = 6.82$, $MSE = 1.91$, $\eta2 = .06$	0.154	[0.150, 0.158]	0.057	[0.050, 0.066]	–0.383	[–0.392, –0.374]
Contemplation Stage	$F = 3.77$, $MSE = 0.91$, $\eta2 = .03$	0.014	[0.005, 0.023]	0.045	[0.027, 0.063]	0.382	[0.363, 0.403]
Action Stage	$F = 20.07$, $MSE = 4.99$, $\eta2 = .15$	–0.079	[–0.089, –0.068]	0.333	[0.312, 0.355]	0.735	[0.711, 0.758]

Notes. NFB = no feedback group; WFB = written feedback group; WOFB = written plus oral collaborative test feedback group. $df1 = 2$; $df2 = 226$; $p < .001$.

lower in the posttest, $\Delta M = -0.066$, 95% CI [–0.072, –0.016].

For the posttest, participants in the three feedback groups significantly differed (a) in their perceived benefit from participating in the TMIC, $F(2, 225) = 23.36$, $MSE = 22.24$, $p < .001$, $\eta^2 = .17$; (b) in their intercultural self-confidence, $F(2, 225) = 5.28$, $MSE = 2.64$, $p < .01$, $\eta^2 = .05$; and (c) in their intercultural self-understanding, $F(2, 225) = 30.01$, $MSE = 19.28$, $p < .001$, $\eta^2 = .21$. As predicted in hypotheses 2a–c, the following contrasts show that the therapeutic effect was highest for individuals who took part in the written plus oral collaborative test feedback: perceived benefit from TMIC, $t(225) = -6.80$, $p < .001$, $d = -0.91$; intercultural self-confidence, $t(225) = -3.25$, $p < .01$, $d = -0.43$; intercultural self-understanding, $t(225) = -7.29$, $p < .001$, $d = -0.97$.

Stages of Change

First, we examined differences between the experimental groups concerning all three stages of change in the pretest. This revealed similar starting values in the precontemplation stage across groups, $F(2, 809) = 0.27$, $MSE = 0.19$, $p = .765$, but different values in the contemplation stage, $F(2, 809) = 16.48$, $MSE = 12.83$, $p < .001$, $\eta^2 = .04$, and action stage, $F(2, 809) = 14.51$, $MSE = 11.62$, $p < .001$, $\eta^2 = .04$. We chose Hochberg's GT2 post hoc test for group comparisons. Concerning the action stage, three subgroups significantly ($p \leq .01$) differed from each other. As displayed in Table 2, highest starting values were found in the written plus oral collaborative test feedback group and lowest in the no feedback group. Two significantly ($p \leq .01$) different subgroups emerged for the contemplation stage with the written plus oral collaborative test feedback group outperforming the written feedback and the no feedback groups.

Second, we computed repeated measures linear mixed models for all three stages of change. In the precontemplation stage, only the interaction effect time × feedback group was significant (see Table 4). Whereas participants in the no feedback and in the written feedback group scored higher on the precontemplation stage in the posttest, values for the written plus oral collaborative test feedback decreased in the posttest. This is in line with hypothesis 3a. For the contemplation stage, values neither significantly increased over time, $F(1, 226) = 6.40$, $MSE = 1.54$, $p = .012$, $\eta^2 = .03$, $\Delta M = 0.147$, 95% CI [0.139, 0.157], nor was the interaction effect (time × feedback group) significant (see Table 4). Hence, hypothesis 3b had to be rejected. Concerning the action stage, a significant main effect was observed, $F(1, 226) = 31.09$, $MSE = 7.73$, $p < .001$, $\eta^2 = .12$. Details about the interaction effects are shown in Table 4; individuals who received written or written plus oral collaborative test feedback reached higher values in the posttest, whereas survey participants without any feedback had lower scores in action in the posttest, which supports hypothesis 3c.

As mentioned before, only for the precontemplation stage were values in the pretest similar across groups. Therefore, mean comparisons in the posttest were computed only for this stage. The experimental groups differed significantly in the precontemplation stage of the posttest, $F(2, 224) = 10.77$, $MSE = 7.05$, $p < .001$, $\eta^2 = .09$. The lowest mean was found for the written plus oral collaborative test feedback group when compared with the written feedback and the no feedback groups, $t(224) = 4.47$, $p < .001$, $d = 0.60$.

Evaluation of the Oral Part of the Collaborative Test Feedback (SHORT)

To examine if the evaluation of the feedback increases after participation in the oral part of the collaborative test feedback (SHORT)

we conducted a paired *t* test. In the pretest, after participants received a written feedback report, the average evaluation mean was 3.78 (*SD* = 0.57). After taking part in SHORT, the evaluation mean significantly increased, *M* = 4.97 (*SD* = 0.70), $t(32) = -9.97$, $p < .001$, $d = -3.52$, which is in line with hypothesis 4a. Furthermore, the average rating of SHORT was high, *M* = 5.24, *SD* = 0.82. The overall evaluation of the feedback significantly correlated with the specific rating of the oral collaborative test feedback SHORT in the posttest, $r(33) = .82$, $p < .001$. Additionally, the more satisfied individuals were with the written feedback and with the oral session of the collaborative test feedback SHORT, the higher was their TMIC-SA total score in the posttest, $r(33) = .43$, $p < .05$ and $r(33) = .38$, $p < .05$, respectively. This supports hypothesis 4b. Moreover, for the posttest the evaluation of SHORT was highly related to the perceived benefit of taking part in the TMIC, $r(33) = .86$, $p < .001$, moderately related to intercultural self-confidence, $r(33) = .35$, $p < .05$, and bordered on significance for intercultural self-understanding, $r(33) = .33$, $p = .060$.

DISCUSSION

This study contributes to the research on intercultural competence development of students in higher education in showing that collaborative assessment might be applied as a personalized, yet economic, intervention.

Key Findings

Except for the contemplation stage, all hypotheses were supported. However, feedback effects were consistent only for individuals who received written plus oral collaborative test feedback, that is, they scored higher on the TMIC-SA, on all therapeutic benefit variables, and on the action stage in the posttest. Moreover, students' values on the precontemplation stage decreased after they participated in the written plus oral collaborative test feedback process, which shows that the intervention changes an individual's perception of having no need to change his or her problematic behavior. Whereas there was a tendency for values to slightly decrease over time in the no feedback group, effects were somewhat inconsistent for students who received a written feedback report. Written feedback had no effect on the TMIC-SA total score and had a negative effect on the perceived benefit from taking part in the test. A potential reason could be that participants could not personally envision how to use the feedback if an assessor did not guide them. Intercultural self-confidence, intercultural self-understanding, and the action stage were positively influenced by written feedback. An interesting finding was that individuals who received exclusively a written report about their intercultural competences were more unwilling to improve their weaknesses in the posttest than they were in the pretest. This shows that written assessment feedback is not necessarily better than receiving no feedback at all, which is in line with previous findings (Lance & Krishnamurthy, 2003). A possible explanation could be derived from attribution theory (Weiner, 1985). Individuals receiving written feedback could attribute a negative outcome to external factors, such as the test characteristics, or to stable aptitude factors, which they think are unchangeable.

Limitations and Future Directions

Drop-out rates in this study were high. However, our univariate analyses of potential systematic differences between students taking part once and students taking part twice in the study revealed no significant results. In general, outgoing students are an attractive sample for different kinds of studies. Therefore, these students receive a high number of invitations

for various research projects. Together with the fact that participation was voluntary, this might be one explanation for the drop-out rates in this study.

Additionally, this study focused on the benefits of collaborative assessment for students going abroad. Therefore, results cannot be readily generalized to other important populations such as adult expatriates. Also, the gain in intercultural competence is somewhat subjective. TMIC uses self-appraisal to measure intercultural competences. When one wants to know if there is an objective improvement after participating in SHORT, using 360-degree feedback (Ward, 1997) might be a fruitful approach. Hereby, an individual is evaluated by several counterparts (e.g., his or her peers, his or her leaders/professors, etc.).

Also, we still do not know enough about the unique components of collaborative assessment that might foster or hinder positive effects for individuals. This leaves several questions unanswered: Which effect would a graphic representation of results, for example using a radar chart, have? Is there a difference between collaborative test feedback administered in person, on the phone, and through new media? Do effects vary if the assessor is male or female? Nevertheless, results show that, especially when time and/or budget is limited, collaborative assessment might be a way to go beyond pure knowledge-building in intercultural competence development of students and hereby sensitize them for their future encounters abroad.

Correspondence concerning this article should be addressed to Deborah Schnabel, Hector Research Institute of Education Sciences and Psychology, Europastraße 6, 72072 Tuebingen, Germany; deborah.schnabel@gmail.com

REFERENCES

Allen, A., Montgomery, M., Tubman, J., Frazier, L., & Escovar, L. (2003). The effects of assessment feedback on rapport-building and self-enhancement processes. *Journal of Mental Health Counseling, 25*, 165–181. Retrieved from http://essential.metapress.com/content/lw7h84q861dw6ytj/

American Psychological Association. (1986). *Guidelines for computer-based tests and interpretations.* Washington, DC: American Psychological Association.

American Psychological Association. (2010). *Ethical principles of psychologists and code of conduct.* Washington, DC: American Psychological Association.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16*, 397-438. doi:10.1080/10705510903008204

Balcazar, F., Hopkins, B. L., & Suarez, Y. (1986). A critical, objective review of performance feedback. *Applied Psychology*, *66*, 314–323. doi:10.1300/J075v07n03_05

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behaviour* (Vol. 4, pp. 71–81). New York, NY: Academic Press.

Bangert-Drowns, R. L., Kulik, C. L., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, *61*, 213-237. doi:10.3102/00346543061002213

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, *5*(1), 7-75. doi:10.1080/0969595980050102

Brislin, R. W., & Bhawuk, D. P. S. (1999). Cross-cultural training: Research and innovations. In J. Adamopoulos & Y. Kashima (Eds.), *Social psychology and cultural context* (pp. 205–216). Thousand Oaks, CA: SAGE.

Craddick, R. A. (1975). Sharing oneself in the assessment procedure. *Professional Psychology, 6,* 279-282. doi:10.1037/h0078597

Crooks, T. (2001, September). *The validity of formative assessments*. Paper presented at the Sixth British Educational Research Association Annual Conference, Leeds, England.

Deardorff, D. K. (2010). *The SAGE handbook of intercultural competence.* Thousand Oaks, CA: SAGE.

DiClemente, C. C., & Prochaska, J. O. (1998). Toward a comprehensive, transtheoretical model of change: Stages of change and addictive behaviors. In W. R. Miller & N. Heather (Eds.), *Treating addictive behaviors* (2nd ed., pp. 3–24). New York, NY: Plenum Press.

Earley, P. C. (1987). Intercultural training for managers: A comparison of documentary and interpersonal methods. *Academy of Management Journal*, *30*, 685–698. Retrieved from http://www.jstor.org/stable/256155

Earley, P. C., & Ang, S. (2003). *Cultural intelligence: Individual interactions across cultures.* Palo Alto, CA: Stanford University Press.

Fantini, A., & Tirmizi, A. (2006). Exploring and assessing intercultural competence. *World Learning Publications*, *1,* 1–75. Retrieved from http://digitalcollections.sit.edu/worlslearning_publications/1

Fecht, J., Heidenreich, T., Hoyer, J., Lauterbach, W., & Schneider, R. (1998). Veränderungsstadien bei stationärer Alkoholentwöhungsbehandlung—Probleme der diagnostik [Stages of change with stationary alcohol withdrawal—Problems in diagnostics]. *Verhaltenstherapie und Psychosoziale Praxis, 30,* 403-419. Retrieved from http://www.dgvt-verlag.de

Finn, S. E. (1996). *Manual for using the MMPI–2 as a therapeutic intervention*. Minneapolis, MN: University of Minnesota Press.

Finn, S. E. (2007). *In our clients' shoes: Theory and techniques of therapeutic assessment*. Mahwah, NJ: Earlbaum.

Finn, S. E., & Kamphuis, J. H. (2006). Therapeutic assessment with the MMPI–2. In J. N. Butcher (Ed.), *MMPI–2: A practitioners guide* (pp. 165-191). Washington, DC: APA Books.

Finn, S. E., & Tonsager, M. E. (1992). Therapeutic effects of providing MMPI–2 test feedback to college students awaiting therapy. *Psychological Assessment, 4*, 278-287. doi:10.1037//1040-3590.4.3.278

Finn, S. E., & Tonsager, M. E. (1997). Information-gathering and therapeutic models of assessment: Complementary paradigms. *Psychological Assessment, 9,* 374-385. doi:10.1037/1040-3590.9.4.374

Fischer, C. T. (1994). *Individualizing psychological assessment.* Mahwah, NJ: Erlbaum.

Fischer, C. T. (2000). Collaborative, individualized assessment. *Journal of Personality Assessment, 74*, 2-14. doi:10.1207/S15327752JPA740102

Fischer, C. T., & Finn, S. E. (2008). Developing the life meaning of psychological test data: Collaborative and therapeutic approaches. In R. P. Archer & S. R. Smith (Eds.), *Personality Assessment* (pp. 379-404). New York, NY: Routledge.

Folds, J. H., & Gazda, G. M. (1966). A comparison of the effectiveness and efficiency of three methods of test interpretation. *Journal of Counseling Psychology, 13*, 318-324. doi:10.1037/h0023708

Gabrenya, W. K., Griffith, R. L., Moukarzel, R. G., Pomerance, M. H., & Reid, P. A. (2012). Theoretical and practical advances in the assessment of cross-cultural competence. In D. D. Schmorrow & D. M. Nicholson (Eds.), *Advances in design for cross-cultural activities* (pp. 317-332). Boca Raton, FL: Taylor Francis.

Hannöver, W., Rumpf, H.-J., Meyer, C., Hapke, U., & John, U. (2001). Der Fragebogen zur Änderungsbereitschaft bei Alkoholkonsum (RCQ-D) [Readiness to Change Questionnaire for Consuming Alcohol (RCQ-D)]. In A. Glöckner-Rist, F. Rist, & H. Küfner (Eds.), *Elektronisches handbuch zu erhebungsinstrumenten im suchtbereich (EHES) [Electronical handbook for assessment instruments for the addictions (EHES)]* (Version 1.00). Mannheim, Germany: Zentrum für Umfragen, Methoden und Analysen (ZUMA).

Hathaway, S. R., & McKinley, J. C. (1942). *Minnesota Multiphasic Personality Inventory.* Minneapolis, MN: University of Minnesota Press.

Hattie, J. A. (1999, June). *Influences on student learning* (Inaugural professorial address, University of Auckland, New Zealand). Retrieved from https://cdn.auckland.ac.nz/assets/education/hattie/docs/influences-on-student-learning.pdf

Hattie, J. A., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*, 81-112. doi:10.3102/003465430298487

Hammer, M. R., Bennett, M. J., & Wiseman, R. (2003). Measuring intercultural sensitivity: The Intercultural Development Inventory. *International Journal of Intercultural Relations, 27,* 421-443. doi:10.1016/s0147-1767(03)00032-4

Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1987). The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist*, *42*, 963-974. doi:10.1037/0003-066X.42.11.963

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6,* 65-70. Retrieved from http://www.jstor.org/stable/4615733

Holmes, J. E. (1964). The presentation of test information to college freshman. *Journal of Counseling Psychology, 11*, 54-58. doi:10.1037/h0046594

Hoyer, J., Heidenreich, T., Fecht, J., Lauterbach, W., & Schneider, R. (2003). Stadien der veränderung in der stationären alkoholentwöhnungstherapie [Stages of change with stationary alcohol withdrawal]. *Verhaltenstherapie, 13*, 31-38. Retrieved from http://psylux.psych.tu-dresden.de/i2/klinische/studium/literatur/195m.pdf

Kelley, C., & Meyers, J. (1995). *The Cross-Cultural Adaptability Inventory (manual).* Minneapolis, MN: National Computer Systems.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*, 254-284. doi:10.1037/0033-2909.119.2.254

Koester, J., & Olebe, M. (1988). The Behavioral Assessment Scale for Intercultural Communication Effectiveness. *International Journal of Intercultural Relations, 12*, 233-246. doi:10.1016/0147-1767(88)90017-X

Lance, B. R., & Krishnamurthy, R. (2003, March). *A comparison of three modes of MMPI–2 test feedback*. Paper presented at the annual meeting of the Society for Personality Assessment, San Francisco, CA.

Landis, D. E., & Bhagat, R. S. (1996). *Handbook of intercultural training*. Thousand Oaks, CA: SAGE.

Leung, K., Ang, S., & Tan, M. L. (2014). Intercultural competence. *Annual Review of Organizational Psychology and Organizational Behavior, 1*, 489-519. doi:10.1146/annurev-orgpsych–031413-091229

Locke, E. A., & Latham, G. P. (1984). *Goal setting: A motivational technique that works*. Englewood Cliffs, NJ: Prentice Hall.

Mendenhall, M. E., Stahl, G. K., Ehnert, I., Oddou, G., Osland, J. S., & Kühlmann, T. M. (2004). Evaluation studies of cross-cultural training programs. In D. Landis, J. M. Bennett, & M. J. Bennett (Eds.), *Handbook of intercultural training* (pp. 129-143). Thousand Oaks, CA: SAGE.

Miller, W. R., & Rollnick, S. (2002). *Motivational interviewing. Preparing people for change*. New York, NY: Guilford Press.

Moin, A. K. (1986). *Relative effectiveness of various techniques of calculus instruction: A meta-analysis* (Unpublished doctoral dissertation). University of Syracuse, Syracuse, NY.

Mosak, H. H., & Gushurst, R. S. (1972). Some therapeutic uses of psychological testing. *American Journal of Psychotherapy, 26,* 539-546. Retrieved from http://psycnet.apa.org/psycinfo/1973-07580-001

Newman, M. L., & Greenway, P. (1997). Therapeutic effects of providing MMPI–2 test feedback to clients at a university counseling service. *Psychological Assessment, 9,* 122-131. doi:10.1037/1040-3590.9.2.122

Poston, J. M., & Hanson, W. E. (2010). Meta-analysis of psychological assessment as a therapeutic intervention. *Psychological Assessment, 22*, 203-212. doi:10.1037/a0018679

Prochaska, J. O., & C. C. DiClemente (1982). Transtheoretical therapy: Toward a more integrative model of change. *Psychotherapy: Theory, Research and Practice*, *19*, 276-288. doi:10.1037/h0088437

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change. Applications to addictive behaviors. *American Psychologist, 47*, 1102-1114. doi:10.1037/0003-066X.47.9.1102

Rosenberg, M. (1979). *Conceiving the self.* New York, NY: Basic Books.

Rosinski, P., & Abbott, G. (2006). Intercultural coaching. In J. Passmore (Ed.), *Excellence in coaching. The industry guide* (pp. 153-169). London, England: Kogan Page.

Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119-144. doi:10.1007/BF00117714

Schnabel, D., & Kelava, A. (2013, July). *Development and validation of the German Test to Measure Intercultural Competence (TMIC): A combined-method approach.* Paper presented at the 12th European Conference on Psychological Assessment, San Sebastian, Spain.

Schnabel, D., Kelava, A., Seifert, L., & Kuhlbrodt, B. (2015). Konstruktion und validierung eines multimethodalen berufsbezogenen Tests zur Messung Interkultureller Kompetenz [Development and validation of a job-related multimethod Test to Measure Intercultural Competence]. *Diagnostica*, *61*, 3-21. doi:10.1026/0012-1924/a000110

Schnabel, D., Kelava, A., van de Vijver, F. J. R., & Seifert, L. (2015). Examining psychometric properties, measurement invariance, and construct validity of a short version of the Test to Measure Intercultural Competence (TMIC-S) in Germany and Brazil. *International Journal of Intercultural Relations, 49*, 137-155. doi:10.1016/j.ijintrel.2015.08.002

Sedikides, C., & Gregg, A. P. (2008). Self-enhancement: Food for thought. *Perspectives on Psychological Science, 3*, 102-116. doi:10.1111/j.1745-6916.2008.00068.x

Straub, J., Nothnagel, S., & Weidemann, A. (2010). *Wie lernt man interkulturelle kompetenz? Theorien, methoden und praxis in der hochschulbildung. Ein handbuch*. [How can intercultural competence be learned? Theories, methods, and practice in higher education. A handbook]. Bielefeld, Germany: Transcript.

Swann, W. B., Jr., & Read, S. J. (1981). Self-verification processes: How we sustain our self-conceptions. *Journal of Experimental Social Psychology, 17*, 351-372. doi:10.1016/0022-1031(81)90043-3

Tafarodi, R. W., & Swann, W. B., Jr. (1995). Self-liking and self-competence as dimensions of global self-esteem: Initial validation of a measure. *Journal of Personality Assessment, 65,* 322-342. doi:10.1207/s15327752jpa6502_8

Tafarodi, R. W., & Swann, W. B., Jr. (2001). Two-dimensional self-esteem: Theory and measurement. *Personality and Individual Differences, 31*, 653-673. doi:10.1016/S0191-8869(00)00169-0

Van Dyne, L., Ang, S., & Koh, C. (2008). Development and validation of the CQS. In S. Ang & L. Van Dyne (Eds.), *Handbook of cultural intelligence: Theory, measurement and application* (pp. 16-38). Armonk, NY: M. E. Sharpe.

Van Oudenhoven, J. P., & Van der Zee, K. I. (2002). Predicting multicultural effectiveness of international students: The Multicultural Personality Questionnaire. *International Journal of Intercultural Relations, 26*, 679-694. doi:10.1016/S0147-1767(02)00041-X

Vulpe, T. (2004, November/December). Training for intercultural competence: Avoiding failures in international assignments. *CERC News, 132*, 1-2. Retrieved from http://www.international.gc.ca/cfsi-icse/cil-cai/pdf/tic-aci-eng.pdf

Waiswol, N. M. (1995). Projective techniques as psychotherapy. *American Journal of Psychotherapy*, *49*, 244-259. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7677204

Ward, P. (1997). *360-degree feedback*. London, England: CIPD Publishing.

Weiner, B. (1985). An attribution theory of achievement motivation and emotion. *Psychological Review, 92*, 548-573. doi:10.1037/0033-295X.92.4.54